MARGINAL MEDIATION: A NEW FRAMEWORK FOR INTERPRETABLE MEDIATED EFFECTS

by

Tyson S. Barrett

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in the subject of

Quantitative Psychology

UTAH STATE UNIVERSITY

Logan, Utah

May 2018

Dr. Ginger Lockhart Major Professor	Dr. Jamison Fargo Committee Member
Dr. Rick Cruz Committee Member	Dr. Michael Levin Committee Member
Dr. Adele Cutler Committee Member	Dr. Mark McLellan Vice President for Research and Dean of the School of Graduate Studies

Contents

Listing of tables

Listing of figures

CHAPTER 1
INTRODUCTION

For precept must be upon precept, precept upon precept; line upon line, line upon line; here a little, and there a little. — Isaiah 28:10, King James Bible

Reproducibility and Applicability

Recently, researchers across the health, behavioral, and social sciences have become increasingly concerned with the reproducibility of research. The concern ranges from asserting that "most claimed research findings are false" (Ioannidis, 2005, pg. 696) to "we need to make substantial changes to how we conduct research," (Cumming, 2014, abstract). Some have come to refer to the situation as a "reproducibility crisis" (Begley & Ioannidis, 2015; Munafò et al., 2017; Taylor & Tibshirani, 2015). Regardless of whether the "crisis" classification is accurate, the evidence is troublesome. For example:

1. The Open Science Collaboration (2015) found that, among 100 replications of articles in three psychology journals, fewer than half were able to be reproduced adequately. Possibly more concerning is that, on average, the replicated effect sizes were about half the size of the original effect sizes.
2. Chang & Li (2015) attempted "to replicate 67 papers published in 13 well-regarded economics journals," (abstract) where only 49% were able to be replicated (even with collaboration from the original papers' authors).
3. Prinz, Schlange, & Asadullah (2011) found that 43 (65%) of the 70 medical research articles assessed had inconsistencies.
4. Begley & Ioannidis (2015) highlight a number of studies that could not be adequately reproduced in the biological and health sciences (see Table 1 of Begley & Ioannidis, 2015).

Some of this is expected (Patil, Peng, & Leek, 2016), especially for those using independent data, for the single replication of the study is not necessarily the "truth," nor is some discrepancy between the original and replication unexpected. Indeed, Patil and colleagues state that several independent replications are needed to draw any concrete conclusions about the reproducibility of any set of results. But even in this more proper context, the low reproducibility levels are worse than expected.

This chapter presents the problems associated with reproducibility and, relatedly, applicability of research. Beyond the inherent importance of reproducibility and applicability, these issues have direct bearing in the prevention sciences—particularly as it comes to mediation analysis and related methodologies—as some of the suggestions for improving cannot be implemented with current approaches. Because mediation analysis allows researchers to assess *how* one variable affects another (e.g., assess the pathways or processes of the effect; MacKinnon, 2008), it is used extensively throughout the field.

However, when results from a mediation analysis cannot be reproduced, interventions built on it are likely to fail with wasted effort in ill-informed areas. To help in this regard, this proposal is designed to develop, evaluate and apply a new approach to mediation analysis that can help improve the reproducibility and applicability of the results.

Definitions of Reproducibility and Applicability

The term “reproducibility” (or “replicability”) is defined in various ways (Goodman, Fanelli, & Ioannidis, 2016). For example, the National Science Foundation defines reproducibility as the ability to take the same data and same methods and produce the same output, whereas replicability is the ability for the results to be replicated with independent data. However, other fields define it precisely the opposite: with replicability about using the same data and methods. Coupled with the fact that both are somewhat synonymous, this has caused confusion on what is meant by various policies aimed at increasing one or the other. To alleviate this confusion, Goodman et al. (2016) created a “New Lexicon” in reference to these constructs. They use the following useful definitions:

1. *Methods Reproducibility* “refers to the provision of enough detail about study procedures and data so the same procedures could … be exactly repeated” (pg. 2) with the same data,
2. *Results Reproducibility* “refers to obtaining the same results from the conduct of an independent study whose procedures are as closely matched to the original experiment as possible” (pg. 2-3) with independent data, and
3. *Inferential Reproducibility* “refers to the drawing of qualitatively similar conclusions from either an independent replication of a study or a reanalysis of the original study” (pg. 4).

Methods reproducibility, then, requires an “open” science framework where researchers release their analysis code, their data, and any measurement instruments to the public (or at least make it available to other researchers upon request). Many journals have already began requiring elements of methods reproducibility (e.g., *Psychological Science*). In contrast, results reproducibility requires meaningful measures of effect size, proper application of statistical methods, and avoidance of bias. When discussing “reproducibility” in the literature, this is often the type that is being referenced. Indeed, unless otherwise specified, the term “reproducibility” refers to results reproducibility in this project.

Notably Goodman et al. (2016) suggest that the third—inferential reproducibility—is the most important of the three areas. The authors argue that this is due to the conclusions often being the most meaningful aspect of a study, especially in the health, behavioral, and social sciences. For example, the *exact* effect size of an intervention may not be as important as the direction and the evidence that it produces a “meaningful” effect. In this way, inferential reproducibility suggests that a finding is reproduced if the same overall conclusions are drawn about an effect. Regardless of relative importance, these three constructs are largely beneficial in the discussions to follow.

Finally, to be applicable, the results not only need to be reproducible as defined above (methods, results, or inferential) but also requires the use of interpretable and actionable results. That is, the result needs to represent reality and needs to be presented in a meaningful form that allows other researchers, laypersons, clinicians, and lawmakers to apply the findings.

Improving with Effect Sizes and Confidence Intervals

Several researchers (Cumming, 2014; Ioannidis, 2005; Munafò et al., 2017) have discussed why the reproducibility of science seems to be low and have made recommendations to improve. Some of these problems are difficult to alleviate and will take time to improve (e.g., reduce conflicts of interest, increase quantitative training). Yet, there are measures that can be taken in the short-term that can help improve reproducibility and applicability of research without major adjustments or costs. Among the most important of these, Cumming (2014) convincingly stated that researchers need to rely far less on Null Hypothesis Significance Testing (NHST; the use of a p-value cut off point [$p<.05$ in many cases]).[1] In its place Cumming (2014) recommends using a more interpretable, yet related construct—effect sizes with confidence intervals. This is a useful strategy to "adopt estimation thinking and avoid dichotomous thinking," (Cumming, 2014, pg. 8), for this dichotomous thinking resulting from NHST may be a driving force in the reproducibility and applicability problem.[2]

For effect sizes and confidence intervals to improve research the way Cumming (2014) intends, the estimates must be in meaningful units (Preacher & Kelley, 2011). To be meaningful, both the *direction and magnitude* need to relay information that can be quickly connected to observable phenomenon. For example, meaningful units include correlations, probabilities or risk, counts, time, and dollars. If effect sizes and the uncertainty in the estimates are in metrics that are difficult to understand, or even in arbitrary metrics, the meaning of the findings can be obscured or misconstrued. Not only can this hurt reproducibility, but can damage the applicability of research results. Therefore, wherever possible, intuitive effect sizes and confidence intervals should be reported.

Mediation Analysis Lacks Intuitive Effect Sizes

Effect sizes and confidence intervals can be used with many statistical analyses. Unfortunately, effect sizes are not well defined in several situations wherein *mediation analysis* is often applied. Because of its importance to the field, an extension to mediation analysis is proposed to make it more interpretable across a wide variety of data types and situations. The following chapters discuss mediation analysis and its need to be extended to increase interpretability, reproducibility, and applicability and the methods to develop and evaluate the proposed approach.

[1] Arguably, NHST was not designed for such extensive use (e.g., in selecting covariates, assessing model fit, inferring important relationships, making bivariate comparisons, etc.). Indeed, the creator of p-values did not like the use of NHST; a sentiment now openly shared by the American Statistical Association (Goodman, 2016).

[2] Although confidence intervals are based on the same information as p-values, if in a meaningful metric, it relays far more information.

CHAPTER 2
MEDIATION ANALYSIS

The point is that mediation analysis provides information regarding possible mediating mechanisms. — David MacKinnon

Introduction

Researchers in the prevention sciences are often interested in *how* one variable may affect another. For this reason, mediation analysis is a key analytic tool in prevention work (Coie et al., 1993) for it allows researchers to evaluate the processes or pathways of an effect—of interventions, risk-factors, and protective-factors alike (Fairchild & MacKinnon, 2009; Hayes, 2009; Iacobucci, 2008; MacKinnon, 2008; MacKinnon, Fairchild, & Fritz, 2007; Shrout & Bolger, 2002). It uses predictors, mediators, and outcomes within a single conceptual model where "the independent [predictor] variable influences the mediator, which in turn exerts an influence on the dependent [outcome] variable," (Serang, Jacobucci, Brimhall, & Grimm, 2017, pg. 1). Thus, it provides an approach to understand the potential process underlying an effect between a predictor and an outcome.

Mediation analysis, as built on linear regression (Edwards & Lambert, 2007; Hayes, 2009), combines two or more regression models to estimate the full conceptual mediation model. It is sometimes referred to as *Conditional Process Analysis* when combined with moderation (i.e., interaction effects; Hayes, 2013). Confirmatory analysis within mediation is well established for a variety of situations (e.g., Lockhart, Mackinnon, & Ohlrich, 2011) while exploratory analysis is beginning to take shape (Serang et al., 2017). Confirmatory mediation has been applied often in health behavior research—showing pathways leading to health-risk behavior such as drug use (Lockhart et al., 2017; Luk, Wang, & Simons-Morton, 2010; Shih et al., 2010; Wang, Simons-Morton, Farhart, & Luk, 2009), tobacco use (Ennett et al., 2001), and alcohol use (Catanzaro & Laurent, 2004).

Knowing the pathway of effect allows clinicians, interventionists and policymakers to target modifiable parts of the pathway. For example, there is evidence that bully victimization in adolescence increases depression, which subsequently increases drug use (Luk et al., 2010). In this example, assuming no confounding, there are at least two immediate targets of intervention: the victimization and the depression. Interventions based on a model without the mediator will be incomplete and may fail to alleviate the risk-factor(s). Further, without the mediating effect included in the model, we are at risk of confounding, causing our estimates to be misleading.

In its simplest form, as shown in Figure 2.1, X is the predictor, M is the mediator, and Y is

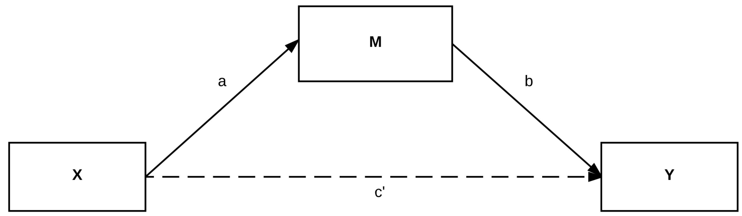

Figure 2.1: Path diagram of a simple mediation analysis model with a single predictor, a single mediator, and a single outcome.

the outcome. The paths labeled a and b make up the mediated effect (i.e., "indirect" effect) of X on Y whereas path c' is the direct effect of X on Y (Hayes, 2009). It is important to note that mediation analysis can become much more complex than that in the figure. Covariates and moderation can be added to the model as well, potentially for a more causal interpretation (Hayes, 2013; Small, 2013).

Definitions

Before discussing mediation further, it is helpful to note some terminology that are often used in the field. In Figure 2.1, X is a predictor or an exogenous variable (i.e., a variable that is not predicted or influenced by something in the model; "independent variable") while M and Y are mediators and outcomes, respectively. These are also known as endogenous variables (i.e., variables that are, in part, predicted or influenced by other variables in the model). These distinctions are particularly useful as it comes to the reproducibility, interpretation and the assumptions of the models.

Reproducibility and Applicability of Mediated Effects

A notable aspect of mediation is that only the predictors can be randomized (i.e., the mediator cannot be randomized in most situations). That is, even when the a and c' paths can portray an experimental manipulation, the b path(s) cannot. Therefore, the need for proper covariates, interpretation, reporting, and replication is even more important in mediation analysis. David MacKinnon said it well: "It is not likely that a true mechanism can be demonstrated in one statistical analysis…. These analyses inform the next experiment that provides more information …" (MacKinnon, 2008, pg. 67).

Therefore, the conceptual model must be considered carefully in light of theory, prior literature, and proper covariates (Lockhart et al., 2011). Even when done properly, replication of mediated effects is important (MacKinnon, 2008). Iacobucci (2008) also recommends (several times) to evaluate competing models and theories—thus presenting the effects and paths in light of alternative model specifications. Further, the interpretation—comprising the magnitude and direction of the effect—needs to be reported with the proper uncertainty and include information on a) bivariate correlations, b) information on all relevant paths (even non-significant ones), c) must include information on the process of variable and covariate selection, d) report standardized and unstandardized results, and e) provide unidentifiable data and code [if possible]. It is important to note that much of this information can be included as supplemental material. In this way, results are reported that can be combined with others in order to provide "convincing evidence of [or lack of] a mediating mechanism", (MacKinnon, 2008, pg. 67).

Additionally, the focus on significance often found in the mediation analysis literature (e.g., Iacobucci, 2008) is problematic to reproducibility and applicability. That is, if the focus is on significance, the effect sizes with the associated uncertainty are often overlooked. This reduces applicable interpretation and relies too heavily on NHST.

Complexity of the Models

One more concept is particularly important for both reproducibility and applicability of research results in regards to mediation analysis. As mentioned previously, the interpretation must be intuitive and straightforward. This does not imply that the model must be simple or straightforward; rather, the information gleaned from it must be. Obtaining informative, actionable results do not require a model to be simple or complex. That is, the theory, empirical work, and the data should be allowed to drive the complexity of the modeling procedure.

This flexibility, however, may be a driver of reproducibility problems. It has been referred to as "researcher degrees of freedom," (Fidler et al., 2016; Simmons, Nelson, & Simonsohn, 2011) wherein a researcher can use many different techniques, designs, and specifications to arrive at their desired significance level. Because of this, Fidler et al. (2016) suggests to use a pre-registry of study design, use sensitivity analysis to show how researcher decisions affected conclusions, and share data and code if possible. Additionally, the simplest, most appropriate modeling procedure/specification should be used.

Frameworks

Two highly related frameworks exist to perform mediation analysis (Iacobucci, 2008). First, as mentioned previously, mediation analysis can be built on linear regression including ordinary least squares (OLS) and generalized linear modeling (GLM; Hayes, 2009, 2013). This requires separate models for the a paths and the b and c' paths, fit independently, to be combined into one mediation model. This approach is flexible in terms of the types of variables and model specifications as compared to the other—structural equation modeling (SEM). For example, performing moderated mediation is straightforward in this framework (Edwards & Lambert, 2007; Hayes, 2013). Ultimately, the regression-based framework is what this proposal builds upon.

Under the SEM paradigm, all the paths are simultaneously estimated, sometimes providing more statistical power (Iacobucci, 2008).[1] This approach notably allows more testing of the full model fit and

[1] The idea that SEM is "superior" to the regression paradigm was refuted by Hayes (2013) by noting that in most

can easily include latent variables but assumes, in general, that all variables are continuous with a multivariate normal distribution. This is a strict assumption that is difficult to assess. However, it has extensions allowing for categorical (generally ordinal) variables to be included, although this changes the estimation procedure. Issues relating to categorical mediators/outcomes are discussed in the "Categorical Mediators and/or Outcomes" sections below.

Assumptions

In his 2008 "Introduction to Mediation Analysis" book, MacKinnon (2008) discusses the assumptions[2] of the modeling procedure. Of these primary assumptions, note that there are no major differences from the assumptions of regression analysis.

1. *Correct Function Form.* In general, mediation assumes a linear relationship between predictors and mediators/outcomes. This can be adjusted using transformations or, more pertinently, generalized linear models (e.g., logistic regression). MacKinnon (2008) also points out that it is assumed the relationships are additive; if they are not, then the correct interactions (moderators) need to be included in the model specification. This, in many ways, needs to be driven by theory and prior literature (Lockhart et al., 2011).
2. *No Omitted Influences.* A key to any mediation analysis is that variables that: 1) correlate with both the predictor and the mediator (path a), 2) correlate with both the mediator and the outcome (path b), or 3) correlate with both the predictor and the outcome (path c') are included in the model. A more general form of this assumption has been termed "sequential ignorability," (Imai, Keele, & Tingley, 2010). This more general form includes a sensitivity analysis to assess how important deviations from this assumption are on the conclusions (Imai et al., 2010; Imai, Keele, & Yamamoto, 2010).
3. *Accurate Measurement.* Random measurement error produces attenuated paths (in large sample sizes) and random bias (in small sample sizes) in regression (Loken & Gelman, 2017) and therefore can affect the paths in various ways (e.g., attenuate the b path which can inflate the c' path). When possible, reliable measures and/or proper latent variable modeling should be used for this assumption to be met.
4. *Well-Behaved Residuals.* The residuals are assumed to be random, "have constant variance at each value of the predictor variable" and "residual error terms are uncorrelated across equations" (pg. 55). The last assumption about uncorrelated errors stems from "No Omitted Influences" for, if there are omitted variables in both equations, the error terms are going to correlate. This is one of the few assumptions that can be assessed in most situations.

With the addition of *temporal precedence* (predictor comes before mediator) and *appropriate measurement timing* (the mediator is measured at the appropriate time when the effect of the predictor has

situations differences in the estimation is extremely minor and will not alter the conclusions. This can be seen in the small effect sizes present in Iacobucci et al. (2007). Further, additional assumptions inherent in the SEM approach may not hold, although some are not easily tested (e.g., multivariate normality). With this said, SEM still provides a powerful framework for mediation analysis.

[2]The assumptions described herein are for both the regression and SEM frameworks for mediation, although, as noted above, SEM has a few additional assumptions as well.

occurred), the resulting estimates are asymptotically (i.e., with a large enough sample size) unbiased, allowing proper (causal) inference regarding the effects' magnitude and direction (MacKinnon, 2008). This interpretation, to aid in reproducibility, needs to be highlighted with the associated uncertainty (e.g., confidence intervals) in a meaningful metric.

Other Considerations

Shrout & Bolger (2002) highlight a number of other important considerations in mediation analysis. First, multi-collinearity can produce problems, especially when it occurs between predictors and mediators. It can distort the statistical power of the analysis, potentially producing misleading results. The second consideration is suppression: "Suppression occurs when the indirect effect $a \times b$ has the opposite sign of the direct effect," (pg. 430). This can, if not interpreted correctly, produce confusing estimates (e.g., a positive indirect path and a negative direct path). Finally, Shrout & Bolger (2002) recommend using bootstrapping (also Hayes, 2009, 2013) to understand the variability in the estimates. This is due to the asymmetric distribution of indirect effects (see Figure 6 in Shrout & Bolger, 2002). Bootstrapping uses repeated random sampling of the data with replacement and estimates the model on that sampled data. Generally, between 500 and 10,000 bootstrapped samples are used to get an accurate confidence interval. Bootstrapping produces as accurate (or more accurate) Type-I error rates than other methods. Because of this, bootstrapping plays a major role in this proposal.

Interpretation

In linear models, the interpretation is simple, straightforward and intuitive. The a path coefficient means: "for a one unit change in X there is an associated a units change in the mediator." Likewise, the b path coefficient means: "for a one unit change in M there is an associated b units change in the outcome, controlling for the effect of X." Finally, the c' path is: "for a one unit change in X, controlling for the effect of M, there is an associated change of c' units in the outcome." The indirect effect is $a \times b$; the total effect is $a \times b + c'$.

Two recommendations in interpreting these models are below:

1. Avoid dichotomous thinking (Cumming, 2014) to best understand mediated effects. That is, each element of the mediation (i.e., the indirect, the direct, and total effects but also the individual a, b, and c' paths) needs to be considered without only trying to answer: "Is there a mediated effect?" Otherwise, researchers can lose sight of each element of the mediation. For example, the various a paths may be important on their own (e.g., if the a path effect size is small then maybe the predictor is not a beneficial place to focus an intervention even though the effect is significant). Therefore, understanding a mediated effect is best told through several avenues: the indirect, direct, and total effects; the individual paths; these effects and paths in light of covariates; among others. This suggestion is also best if those effects are in meaningful and intuitive metrics.

2. Once the analysis ventures into non-normal, non-linear relationships, the interpretation becomes more difficult—particularly when it comes to the indirect and total effects. For example, if the mediator is binary, often logistic regression is used to assess the a path. But that changes the a path interpretation to: "for a one unit change in X, there is an associated a log odds units change in the mediator." This interpretation is anything but intuitive. In general, the log odds are transformed into odds ratios, which improve the interpretation. But these units do not mix well with other units. This is detrimental to understanding the indirect and total effects as will be discussed further in the sections below.

Analytic and Interpretation Issues with Mediation Analysis

Categorical Mediators and/or Outcomes: Reproducibility and Applicability

Mediation analysis is more difficult when the mediator and/or the outcome are not continuous, including dichotomous variables (e.g., an individual either uses marijuana or not), ordinal variables (e.g., the self-reported confidence in social settings), other polytomous variables (e.g., classes of delinquency subtypes), and count variables (e.g., number of hospital visits). These data situations are difficult because mediation analysis requires the mediators and/or outcomes to be continuous and approximately normal (to meet the assumption of *well-behaved residuals*).

There are several strategies taken in the literature to address this problem. However, each makes its own set of assumptions and each contains limitations in interpretation. The variability in approaches and the subsequent interpretations make combining results across studies far more difficult—reducing the chance to concretely show relationships via meta-analyses and systematic reviews. The difficulty of these data situations are likely reducing reproducibility and applicability for a number of reasons:

- It may be easier to ignore the assumptions that are violated when using categorical mediators and/or outcomes. Results from these analyses may not be valid.
- Different approaches produce varying assumptions and interpretations. This can be difficult for other researchers, clinicians, lawmakers, and laypersons to keep straight, possibly leading to misunderstandings regarding results and their validity.
- Analyses with categorical outcomes are not typically well-emphasized in graduate training, even in more simple modeling techniques, not to mention more complex techniques like mediation analysis. With fewer individuals well-trained, more errors are likely in analyzing these data.
- The interpretation regarding analyses with categorical outcomes is often less intuitive than with continuous outcomes. This can produce a higher cognitive load for both the researchers and those utilizing the study.

With this in mind, the following subsection discusses the current approaches to mediation analysis with categorical mediators and/or outcomes and the assumptions these approaches make.

Table 2.1: The various appraoches to handling mediation with categorical mediators/outcomes.

Approach	Pros	Cons
1. Series of logistic regressions	Simple to apply in most software	Ignores some information, cannot obtain indirect effect size
2. Use SEM's approach (polychoric correlation)	Powerful, well-designed, Easy to implement with proper software	Only works with ordinal variables, only standardized effect sizes
3. Standardize the coefficients	Provides significance test of indirect effect	Assumptions (distributions), difficult to interpret beyond p-value
4. Interpret each path separately	Simplest approach with proper models	Ignores some information, cannot obtain indirect effect size
5. Pretend all variables are continuous	Simplest approach	Purposeful mis-specification, poor model fit

Categorical Mediators and/or Outcomes: Problems and Current Approaches

> "The quest for sound methods of incorporating categorical variables is perhaps the last dilemma in mediation analysis that lacks a strong solution—it's the 'final frontier,' " (Iacobucci, 2012, pg. 583).

Although there are likely other "frontiers", categorical mediators and/or outcomes provide several challenges. Iacobucci (2008, 2012) thoroughly discusses the issues of assessing categorical variables within a mediation analysis and some of the current practices along with their associated problems[3]. Four approaches are of note here: 1) a series of logistic regression models, 2) polychoric correlation in SEM, 3) the method suggested by Iacobucci (2012) regarding standardization, and 4) interpret each path separately. A fifth, but certainly least, is to ignore the distribution of the outcomes and purposefully misspecify the model. These are highlighted in Table 2.1.

Of these approaches, only the SEM approach allows any proper estimation of effect sizes. This approach was developed by Muthén (1984), which uses polychoric correlations (for ordered categorical variables) within the structural equation modeling framework. In this approach, Muthén (1984) assumes that the ordered categorical manifest variable results from a continuous latent variable, "with observed categorical data arising through a threshold step function," (Iacobucci, 2012, pg. 583). In many cases, a continuous latent variable is a reasonable assumption. For example, in a binary variable describing whether an adolescent has ever smoked marijuana, assuming a continuous latent variable comprised of the probability of smoking may adequately represent reality. This method relies on using a probit regression "to model the relationship between the observed categorical variable and the latent *normally distributed* variable," (MacKinnon, 2008, pg. 320, emphasis added) and requires large sample sizes (Iacobucci, 2008). Using this latent-observed submodel, the probit is used as the threshold value to estimate the full model. Notably, both the assumption of a normally distributed latent variable and the requirement of large sample sizes are limitations of this approach.

[3]For the present, only situations when a mediator or outcome is categorical is under consideration since a categorical predictor poses very few problems (Iacobucci, 2012).

The other approaches, including the series of logistic regression models and the interpretation of each path separately, do not allow any effect size estimation of indirect effects. These estimate each path well, but “the results [cannot be] put together because the resulting ‘path coefficients’ are not easily integrated and interpreted,” (Chapter 5, “Solutions to the Categorical Problem”, paragraph 5). The third approach, recommended by Iacobucci (2012), suggested using a new standardization solution using logistic regression output (for a categorical variable) that can combine linear and logistic models’ estimates. It is built on the same idea as the Sobel test (MacKinnon, 2008; Sobel, 1982) but is more flexible; for “the mechanics of testing for mediation do not [need to] change whether the variables are continuous or categorical or some mix,” (Iacobucci, 2012, pg. 593). However, MacKinnon & Cox (2012) criticized this approach. Ultimately, the most fatal flaw may be that a focus on the significance (relying on Null Hypothesis Significance Testing) of the effects without regard to their effect size in meaningful terms is less reproducible and interpretable.

In the end, none of these approaches can flexibly handle categorical variables of various kinds (e.g., binary, multinomial) or other non-normal distributions (e.g., costs, counts); none can produce intuitive and meaningful effect sizes and confidence intervals across these variable types; and none can consistently combine two differing types of estimates (e.g., binary mediator with continuous outcome, count mediator with binary outcome). Although the current approaches are useful in some situations—particularly the SEM apporach—a more complete framework is needed.

Conclusions

Mediation analysis is a powerful framework for understanding the processes by which one variable influences another. The assumptions are not much more than that of regression analysis. The interpretation, in linear models, is straightforward and simple. However, once the analysis ventures into non-normal, non-linear relationships, the interpretation becomes more difficult—particularly when it comes to the indirect and total effects.

In the end, Iacobucci (2012) is correct in saying this problem “lacks a strong solution” (pg. 583). Although important information can be obtained from the current methods, mediation analysis with categorical mediator(s) and/or outcome(s) still misses the mark on intuitive, meaningful effect sizes without a reliance on p-values.

This proposal aims to alleviate these issues by integrating a post-estimation approach known as Average Marginal Effects (AMEs) within the framework of mediation. The following chapter introduces AMEs, showing their benefit in interpretation and reporting when working with non-normal variables within generalized linear models. In turn, this integration can allow simple and meaningful interpretation across variable types and combinations thus far shown to be problematic.

CHAPTER 3
AVERAGE MARGINAL EFFECTS

Any fool can know. The point is to understand. — Albert Einstein

Introduction

When the outcome variable is not continuous and/or has a distribution far from normal, researchers in prevention science generally use a generalized linear model (GLM). The power of GLMs is clear when you consider the broad range of situations it estimates with asymptotic consistency.[1] The problem with GLMs is that the estimates are not in an easily interpretable form. For example, in logistic regression (one type of GLM), the results are in "log-odds". To overcome this lack of interpretability, a simple exponentiation of the coefficient produces what is known as an odds ratio. Similarly, Poisson regression (another form of GLM), with an exponentiation, produces the risk ratio. Although some fields have adopted odds ratios (or relative risk, risk ratios, and other related metrics), these metrics have notable shortcomings.

1. Most can be difficult to understand (i.e., many are not intuitive).
2. They cannot be combined with other metrics in a meaningful manner.

The second is particularly important in the case of mediation analysis given the importance of the indirect effect (the combination of the a and b paths). If a is not in a unit that can be combined with b, then obtaining a meaningful indirect effect is not generally possible.

Problems with Odds Ratios and Others

Data have suggested that individuals, although with some variability, are able to intuitively grasp the meaning of phrases such as "highly probable" or "not likely" (Heuer Jr., 1999). Yet, this same intuition is not found in odds or odds ratios. For example, Montreuil, Bendavid, & Brophy (2005) found, of 84 articles in several epidemiology journals that used odds ratios, only 7 (8.3%) accurately interpreted the odds ratio and 22 (26%) interpreted odds as though they were probabilities ("risk"). Figure 3.1 highlights the large discrepancy between the odds ratio and the risk. Ultimately, reporting odds and interpreting them as risk is common. Given its commonality and the limitation that it cannot be combined with other metrics, it is time to consider other strategies.

[1] Asymptotic consistency refers to the ability to, as the sample size increases, produce estimates that converge to the proper value.

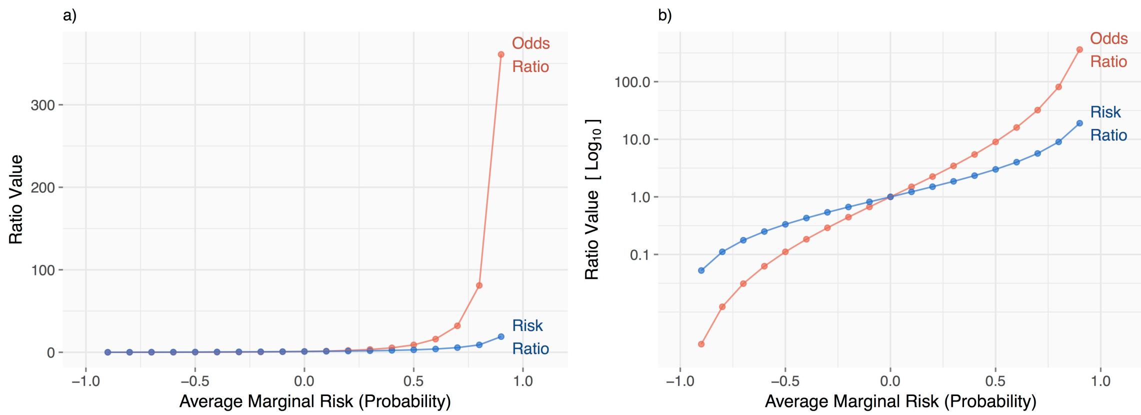

Figure 3.1: a) Comparison of odds ratio and risk ratio with the average marginal risk (probability). b) Same comparison as a) but the y-axis is rescaled (log_{10}) to better show the negative marginal risk comparisons. Both highlight the discrepancy between odds and risk ratios at various levels of marginal risk and that neither approximate the marginal risk.

Additive vs. Multiplicative Interpretations

In most quantitative research designs, the investigators are seeking information on the average effect in a population, whether this refers to an average difference across groups or an average change in the outcome for a given change in the predictor. Generally speaking, the average effect is referring to the marginal effect (i.e., the effect of a small change in the predictor in the outcome's units). In the linear regression framework, the average effect is the estimated coefficient and is interpreted additively—a one unit change in the predictor is associated with an X unit change in the outcome. Conversely, outcomes such as OR are multiplicative. Being multiplicative changes the interpretation to: a one unit change in the predictor is associated with an X times change in the outcome. Although subtle, the difference is important, especially for multi-part models (e.g., mediation analysis).

Being multiplicative indicates that the effect of the predictor changes based on the level of the predictor. For example, if the predictor is high, a small change in the predictor may have a big effect while if the predictor is low, a small change in the predictor has little effect. Figure 3.2 shows this phenomenon, where, in the outcomes original units there is an exponential function. In part a) of the figure, it is clear that a change from 2 to 3 in the predictor has a much larger effect than a change from 0 to 1. A regression would not work well here. If a log transformation is used, the relationship would be linear (and a regression can be used) but the interpretation becomes multiplicative.

Additive interpretations are generally the most intuitive and require less cognitive resources to understand the pattern being portrayed. In a multiplicative framework, simplicity in understanding the effect intuitively is somewhat lost (Iacobucci, 2008). Among others, this is one reason why Stata provides the `margins` command when dealing with two-part hurdle models.[2] These models break up the model-

[2]These models are often used for zero-inflated count outcomes.

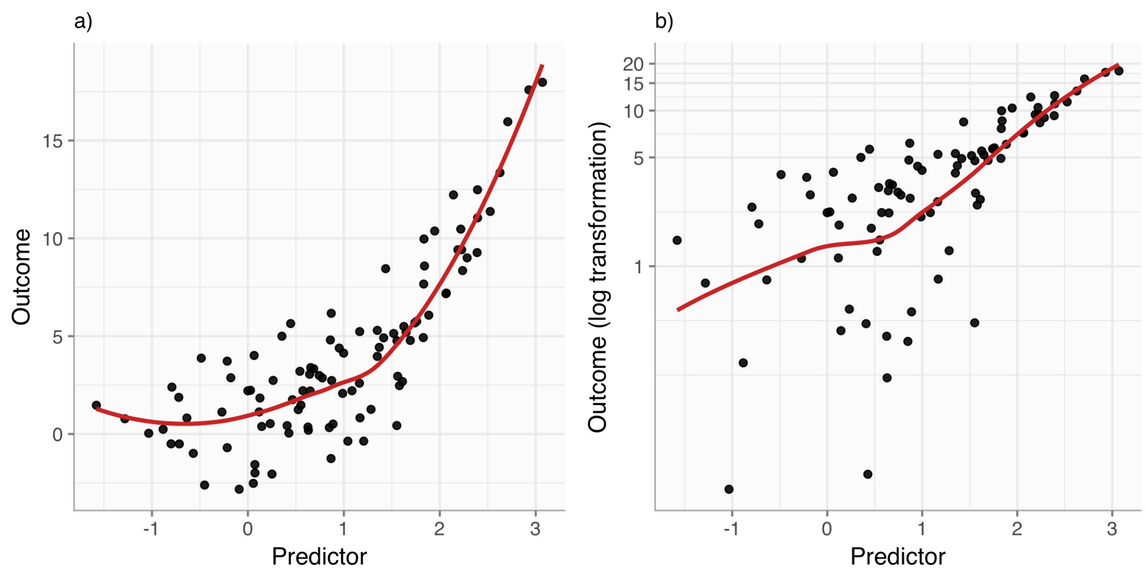

Figure 3.2: Demonstration of a non-linear relationship. a) The outcome is an exponential function of the predictor. b) When log transforming the outcome, the relationship becomes fairly linear. So the interpretation in the log transformed scale is additive, but once it is put back into the original units, it is multiplicative.

ing into two parts: one for the binary part and one for the count part. In order to combine the two parts, Stata allows a transformation known as the average marginal effect (discussed below) that makes the two parts both additive, and therefore easily combined (as is desired in mediation analysis).

Average Marginal Effect

Why Consider Average Marginal Effects?

When using GLMs, the model is fit with a link function (e.g., "logit", "probit", "log"). This change causes the marginal effect to rely on the values of the covariates in the model. This is well illustrated through an example. Say a logistic regression model was fit to the data, as shown in Equations 3.1 - 3.4, for p predictors.

$$logit(Y_i) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i \tag{3.1}$$

$$log(\frac{Prob(Y_i = 1)}{1 - Prob(Y_i = 1)}) = \beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i \tag{3.2}$$

$$\frac{Prob(Y_i = 1)}{1 - Prob(Y_i = 1)} = e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i} \tag{3.3}$$

$$Prob(Y_i = 1) = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i}}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_i + \epsilon_i}} \tag{3.4}$$

This implies that the marginal effect of, say, X_{i1} is:

$$\frac{\delta Y}{\delta X_1} = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i}}{(1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j X_{ij} + \epsilon_i})^2} \tag{3.5}$$

That is, the effect of the predictor X_1 (i.e., the derivative) *depends* on the level of each covariate (all X_js). This, understandably, complicates the interpretation. To simplify the interpretation, two major approaches have been used to understand the marginal effect.

Two Approaches

First, the marginal effect at the mean is a simple, but often unrealistic, approach. In it, the mean value of each covariate is put into Equation 3.5. This simple approach presents problems: 1) some covariates' means may represent impossible values [e.g., dummy coded race variable could be averaged to be 0.4 which is not a meaningful value] and 2) it represents the effect at the mean but not the average effect.

The AME, in contrast, is the *averaged* marginal effect across all observed values in the data. That is, each individual (in a cross-sectional example) has a marginal effect of a given predictor associated with her set of characteristics (covariates). Using these individual marginal effects, the AME is simply the *average* of the marginal effects. In this way, no "impossible" values are used and it is representative of the effect seen in the sample. In linear models, the AME is the same as the original estimates. This is intuitive given the AME is the marginal effect in the outcome's original units–the exact interpretation of the estimates in a linear model. However, as stated previously, in GLMs the estimates are not in the original units and therefore must be estimated via a post-estimation calculation described below.

Definition of the Average Marginal Effect

In an instructive paper about the (now defunct) routine called "margeff" in Stata, Bartus (2005) highlights how the AME can be calculated—including the mathematical definition—and the benefits of AMEs compared to other related methods. The AME is a post-estimation calculation—it uses the model estimates and the data to provide the average effect. Bartus (2005) provided the definition of this post-estimation procedure of a continuous predictor as:

$$AME_k = \beta_k \frac{1}{n} \sum_{i=1}^{n} f(\beta x_i) \tag{3.6}$$

where f refers to the derivative of the estimate with respect to x_i, the βx_i refers to the linear combination of the predictors, and AME_k is the average marginal effect for the kth variable.

Table 3.1: The generalized linear model link functions with their associated units of interpretation. Note: This list is not exhaustive and there are likely more GLMs that are used within prevention research.

Link Function	Average Marginal Effect
Identity	Original Continuous Unit
Logit	Risk
Probit	Risk
Poisson	Count
Gamma	Original Count Unit
Negative Binomial	Count

Relatedly, the AME of a dummy coded variable is:

$$AME_k = \frac{1}{n}\sum_{i=1}^{n}\big[F(\beta x_i|x_i = 1) - F(\beta x_i|x_i = 0)\big] \tag{3.7}$$

where $F(\beta x_i|x_i = 1)$ is the predicted value of the ith observation when the dummy variable equals one and $F(\beta x_i|x_i = 0)$ is the predicted value when the dummy value equals zero. Both the continuous and dummy variable AME are the averages (as the name implies)—averages of the effect of a one unit increase of the x_i variable.

Confidence Intervals

In general, two approaches are taken to estimate the confidence intervals of AMEs. The first approach is the delta method, which provides standard errors (StataCorp, 2015). Although beneficial, the second—bootstrapped confidence intervals—have proven accurate for both the AME and mediation analysis. Therefore, this proposal uses bootstrapped confidence intervals as described in Chapter 2.

Interpretation

Table 3.1 presents the various units that the AME will produce for the various GLMs. It is important to note, that although specifically stated only for some, all AMEs are in their original metrics whether they be probabilities, counts, or something else. The interpretation, then, is "for a one unit change in the predictor there is an associated [AME] change in the outcome."

Benefits and Limitations

There are several benefits of this framework.

- *Intuitive Interpretation and Few Assumptions.* The first, and most obvious, benefit to using AMEs is the simplicity of the interpretation. The effect is in the units used in the modeling; it is additive (i.e., the effect is the added increase or subtracted decrease of the outcome); it provides an interpretation that imitates that of ordinary least squares regression. Relatedly, there are no difficult modeling assumptions directly tied to AMEs. Instead, the underlying models' assumptions that are used to get the AME is what is important. The only additional assumption with AME is that the effect is linear enough to be represented by an additive value (after accounting for the covariates).

- *More Generalizable and Robust.* There is evidence suggesting that AMEs are more robust to problems associated with GLMs (including logistic regression) such as unobserved heterogeneity and model mis-specification (Mood, 2010; Norton, 2012). This allows the estimates to be more generalizable to individuals outside of the sample.
- *Low Computational Burden.* Given AMEs are a post-estimation calculation, no new models need to be fit. Instead, using the estimates of the models, the average marginal effects can be calculated. The most burdensome of the calculation is the bootstrapped confidence intervals.
- *Broadly Applicable.* The AME applies to any of the generalized linear models including logistic, Poisson, gamma, beta, negative binomial, and two-part hurdle models. This provides extensive flexibility in modeling, and, once applied to mediation, will allow flexibility in modeling based on the correct functional form.
- *Two-Part Models.* Particularly pertinent to this proposal is that the calculation of AMEs have been applied to two-part models, generally of the hurdle model types, as stated earlier. In fact, this is a common routine in the Stata statistical software. This provides valuable support for the proposed approach of using AMEs in mediation analysis.

Notably, the interpretation of AMEs hold to the assumption (as found in all regressions) that it is reasonable to adjust a single covariate while holding all others constant. This may not hold in reality, although it may be necessary to gain an understanding of the individual effect of a single variable. In data that are not representative of the population (e.g., non-random sample), AMEs may be biased given an over-representation of certain covariate values may be present. This is an overall modeling problem, since GLMs also assume a random sample. In this way, this problem is not specific to AMEs.

Conclusions

The Average Marginal Effect produces intuitive, interpretable, additive, broadly applicable estimates of effect. They have been applied to two-part models, not unlike mediation analysis, demonstrating their utility in difficult modeling situations. The following chapter discusses the integration of AME with mediation analysis—termed *Marginal Mediation*— with its interpretation and assumptions, its benefits and limitations, and the basic procedures for its use.

CHAPTER 4
MARGINAL MEDIATION

Without an interpretable scale, it is difficult to use effect size to communicate results in a meaningful and useful way. — Preacher and Hayes, 2011

Introduction

The proposed integration of average marginal effects and mediation analysis is designed to resolve two major obstacles currently found in mediation analysis:

1. The difficulty of performing mediation analysis with categorical mediators and/or outcomes, and
2. The lack of reliable and flexible effect size estimates in mediation analysis—especially with categorical mediators and/or outcomes.

These issues are relatively common in prevention work (e.g., B. Hoeppner, Hoeppner, & Abroms, 2017) and the present approaches are not adequate—as was discussed at length in the previous chapters. In this chapter, the integration of Average Marginal Effects (AMEs) and mediation analysis—*Marginal Mediation*—is proposed, including its interpretation and assumptions as well as its benefits and limitations. It is expected that this adjustment to both the modeling and the interpretation will increase both reproducibility and applicability of the results and conclusions.

Effect Sizes

> "First, virtually all effect size indices should be scaled appropriately, given the measurement and the question of interest. Without an *interpretable scale*, it is difficult to use effect size to communicate results in a meaningful and useful way…. Second, it should be emphasized that effect size estimates are themselves sample statistics and thus will almost certainly differ from their corresponding population values. Therefore, it is important to report confidence intervals for effect sizes…" (Preacher & Kelley, 2011, pg. 95, emphasis added).

In this chapter, it is shown that both of these aspects of proper effect size estimation and reporting are adequately represented when using Marginal Mediation. Of particular note, unlike many effect sizes that are only useful in certain research questions, the effect sizes produced by AMEs—and thus found in Marginal Mediation—are flexibly oriented to "be scaled appropriately" to best "communicate results in a meaningful and useful way" for a wide variety of situations. Regarding mediation analysis specifically, Preacher & Kelley (2011) continue: "it is important to develop a way to gauge the effect size of the product term ab itself," (pg. 95). That is, not only does the effect size of the individual paths need to be meaningful but the product of $a \times b$ must be as well. It can be added that a meaningful effect size of the total effect is also necessary (e.g., $a \times b + c'$) in a way that allows $a \times b$ and c' to be comparable.

Definition

The form of the general Marginal Mediation model, including the post-estimation step, are demonstrated in the following equations, where Equations 4.1 and 4.2 demonstrate the mediation estimation while Equations 4.3 and 4.4 show the post-estimation procedures.

$$M_{ij} = a_0 + \sum_{k=1}^{p} a_k x_{ki} + \epsilon_i \quad \text{for j = 1, ..., m mediators} \tag{4.1}$$

$$Y_i = \beta_0 + \sum_{j=1}^{m} b_j M_{ij} + \sum_{k=1}^{p} c_k^{'} x_{ki} + \epsilon_i \tag{4.2}$$

for the ith individual, for $k = 1, ..., p$ predictors, and $j = 1, ..., m$ mediators. The paths are all labeled with their common term (i.e., path a is labeled a). Combining these two equations provides the full mediation model. Using these models, we apply the post-estimation of the average marginal effects as presented by Bartus (2005). For a continuous x variable, the average marginal effect of path a is:

$$AME_k^a = a_k \frac{1}{n} \sum_{i=1}^{n} f(ax_i) \tag{4.3}$$

where f refers to the derivative of the estimate with respect to x_i, ax_i is the linear combination of the predictors, and AME_k^a is the average marginal effect of the a path for the kth variable. Similarly, the AME of a dummy coded variable in the a path is:

$$AME_k^a = \frac{1}{n} \sum_{i=1}^{n} \Big[F(ax_i | x_i = 1) - F(ax_i | x_i = 0) \Big] \tag{4.4}$$

where $F(ax_i|x_i = 1)$ is the predicted value of the ith observation when the dummy variable equals one and $F(ax_i|x_i = 0)$ is the predicted value when the dummy value equals zero. Notably, these same post-estimation equations can be used for the b and c' paths as well.

Interpretation

The interpretation of Marginal Mediation is based on the original units of the mediator(s) and outcome(s). Because there are so many possible combinations of GLM types within mediation analysis, instead of outlining every combination, the basic principles will be presented that applies to all situations.

Principle 1: The individual paths are interpreted based on the corresponding endogenous variable's original metric.

The individual paths have interpretations identical to those of AMEs, as discussed in Chapter 3. Therefore, the a path depends on the type of mediator and modeling approach chosen. For example, for a binary mediator, the a path is a probability (risk); a mediator representing a count has an a path that is in the same count units.

Principle 2: The indirect effect, as a combination of the a and b paths, are interpreted based on the outcome's original metric.

The indirect effect is joining two paths, possibly of different metrics. However, the intuition is simple: the entire effect will be in the outcome's original metric. An example may be beneficial to highlight this principle.

Suppose there are data with a hypothesized binary mediator (depression or no depression) and a continuous outcome (number of hospital visits). A logistic regression is used to model path a and a linear regression is used for the b and c' paths. After calculating the AME of the paths, path a is in units of risk (of depression) while path b is the difference in the number of hospital visits between depressed and not depressed individuals. Combining these two metrics is shown in Figure 4.1.

Another way to understand how these combine in such a straightforward manner is to consider the marginal effects (i.e., the derivatives) of each part of the indirect effect:

$$\frac{dM}{dX} * \frac{dY}{dM} = \frac{1}{k_1}\frac{dY}{dX} \tag{4.5}$$

That is, a change in M for a change in X multiplied by a change in Y for a change in M equals some of the change in Y for a change in X. The derivative, then, is additive—a small change in X results in a small change in Y without depending on the value of the predictor itself. In GLMs, without using the AME, the derivatives do not seamlessly combine. The example below demonstrates if a GLM, sans AME, is used to assess the a path.

$$\frac{dM}{dX} f(X) * \frac{dY}{dM} \neq \frac{dY}{dX} \tag{4.6}$$

where $f(X)$ is an arbitrary function of X (depending on the model), and X is the only predictor. In this case, the small change in X has a differential effect depending on the value of X itself (e.g., Figure 4.1). This overall reasoning holds across GLM types (e.g., logistic, Poisson, negative binomial).

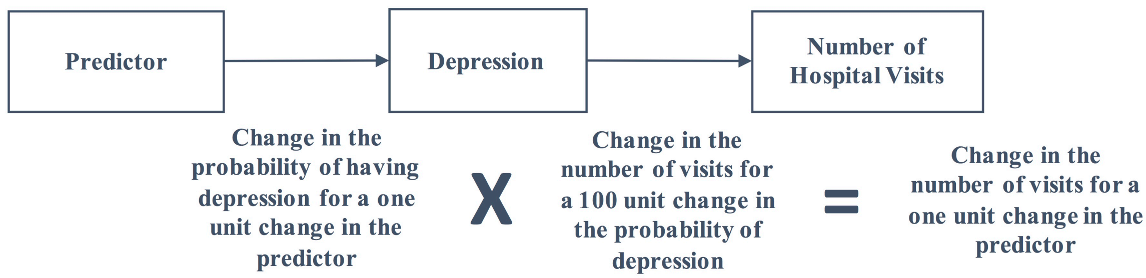

Figure 4.1: Given that path a is additive and in probability units, combining it with the difference in the number of visits for a change in probability is straightforward. Here, $a \times b$ essentially says: "There is an increased probability of a of having depression for a one unit change in the predictor, which if the individual does have depression, that in turn increases the number of hospital visits by b." That is, the change in the number of visits for a one unit change in the predictor indirectly through the mediator.

Principle 3: Both the direct and total effects are interpreted based on the outcome's original metric.

Similar to Principle 2, the direct and total effects are in the outcome's original units. This is intuitive given that the AME of the direct effect is in the outcome's units, and it is not combined with any other path. For the total effect, the indirect and direct effect (which are both in the same units) are added together to get the complete effect. It is expected that, as in linear regression, $a \times b + c' = c$. That is, the indirect and direct effects together will equal the total effect as found in the simple regression $Y = \beta_0 + cX_1$. This expectation will be tested, as described in the next chapter.

Assumptions

The assumptions inherent in Marginal Mediation are the same as those presented in Chapter 2 regarding mediation analysis (see Chapter 2). The only additional assumption regards the ability of the effect to be represented additively (i.e., can the effect be represented linearly after accounting for the marginal effect at each observed level of the covariates?). In linear models, this is already included as an implicit assumption. For other models, although the relationship is not entirely linear, taking the average of the effects across the covariates is assumed to be representative of the relationship.

Reproducibility and Applicability

As mentioned previously, categorical mediators/outcomes are generally not difficult to model using GLMs—only the interpretation is difficult. Generalized linear models, in conjunction with AMEs, will allow researchers to use more correct functional forms, thereby reducing the justification used to fit poorly specified models that have an easier interpretation.

With this framework, Marginal Mediation can be applied across the GLM spectrum and essentially any combination of GLM types. For example, Marginal Mediation is defined when the mediator is binary and the outcome is continuous; when the mediator is a count and the outcome is ordinal; when the mediator is continuous and the outcome is binary. Each has a simple, yet informative, interpretation as outlined by the principles above. This attribute alone can increase the reproducibility of research using mediation analysis.

Additionally, the interpretation across the paths and effects is straightforward and flexible. Other researchers, laypersons, lawmakers, and clinicians can assess the magnitude, the meaning, and the utility of findings much easier—thus, increasing the reach and impact of research. In mediation analysis, this can prove largely beneficial given the already complex nature of the modeling scheme. By simplifying the interpretation, less cognitive resources are required to gain a basic understanding of the findings; instead, more resources can be used to understand how to apply it and assess future research questions based on the findings.

Given the potential benefits of Marginal Mediation, the following chapter discusses the methods proposed to fully develop, evaluate, and apply this new approach.

CHAPTER 5
METHODS

A promise is a cloud; fulfillment is rain. — Arabian Proverb

Introduction

As presented in the previous chapter, Marginal Mediation has the ability to simplify the interpretation of mediated effects in a wide variety of situations, particularly in situations where an interpretation otherwise does not exist (e.g., indirect effects when the mediator or outcome is categorical). In this chapter, methods fashioned to develop Marginal Mediation and evaluate its performance are proposed via three phases:

1. Development of Marginal Mediation
2. Monte Carlo Simulation Study of Marginal Mediation
3. Application of Marginal Mediation

These phases are designed to provide the theory and the software to perform Marginal Mediation, assess the method's ability to accurately estimate the underlying effects, develop the guidelines of its use in finite samples, and apply it to real-world prevention data about the pathways between chronic illness (i.e., chronic migraines and asthma) and delinquency/substance use. Below, each phase is described in depth.

Phase 1: Development of Marginal Mediation

To be useful to public health, psychological, and prevention researchers, the incorporation of average marginal effects within mediation analysis must happen in two ways: in theory and in software. Although the basic framework has been designed, the software and the nuances of the method (e.g., incorporating moderation, assessing the relationship between indirect and total effects, delineating additional model assumptions) need to be developed and assessed.[1] Therefore, this phase is focused on understanding the properties of Marginal Mediation and on developing the software necessary to perform it.

Properties of Marginal Mediation

Building on the mediation framework discussed by Hayes (2009) and by Edwards and Lambert (2007), Marginal Mediation will be established on linear regression—either ordinary least squares for continuous outcomes/mediators or maximum likelihood for categorical outcomes/mediators. In this frame-

[1] It is important to note that when referring to the properties of the estimates we are referring to four distinct, but related, estimates: the a path, the b path, the $a \times b$ (indirect) path, and the c' (direct) path.

work, two or more regression equations are combined to provide the overall mediation model as discussed in Chapter 2. The proposed method will add a post-estimation step into this mediation framework.

The form of the general Marginal Mediation model, including the post-estimation step, were discussed in Chapter 4. Using this framework, various considerations will be made in the development of the method. First, an appropriate manner in which to integrate moderation (interaction effects) into the framework is important. Because of the work by Edwards and Lambert (2007), this will likely include assessing the *reduced form*[2] of the models. Second, it has been noted by MacKinnon (2008) that in non-linear models the $a \times b + c'$ generally does not equal the c path as it does in linear models (Chapter 11). It is expected that these will equal within the Marginal Mediation framework. Third, an assessment of the various assumptions being made across possible model specifications is required for proper understanding of the method. Its robustness to these assumptions will be assessed in Phase 2 (the Monte Carlo simulation study of the finite properties of Marginal Mediation). Finally, these considerations need to be made for the various mediator/outcome types (multinomial, count, ordinal, dichotomous).

Software Development

A major aspect of this first phase is the development of the software for researchers to apply Marginal Mediation. This software will be provided via the R statistical environment given R is free, widely used by researchers in prevention, and extensions to the environment via "packages" are easily disseminated through the Comprehensive R Archive Network (CRAN). The package will be called `MarginalMediation` and will be available for download via CRAN and GitHub. It will consist of a number of functions to fit the model and assess the model's fit while efficiently producing the paths and effects in proper units. A simple diagram displayed in Figure 5.1 demonstrates the mechanics of the main function—`marginal_med()`.

The package will apply the best practices for both computational speed and user readability (Wickham, 2015), allowing other researchers to extend the package more easily. Several built-in tests will inform the functionality of the package before beginning Phase 2. These tests will be performed on Linux, Mac, and Windows platforms. Finally, the package will use Git as the version control software.

Potential Problems and Solutions

The main issue that may come up within this phase is that the software may take longer than expected to develop. The necessary functions will be developed first so that the next phase may continue on schedule. The usability that will be important in the disseminated version will be developed last.

[2]Reduced form refers to only having exogenous variables on the right hand side of a regression equation (i.e., substituting the predictors of the mediators into the equation).

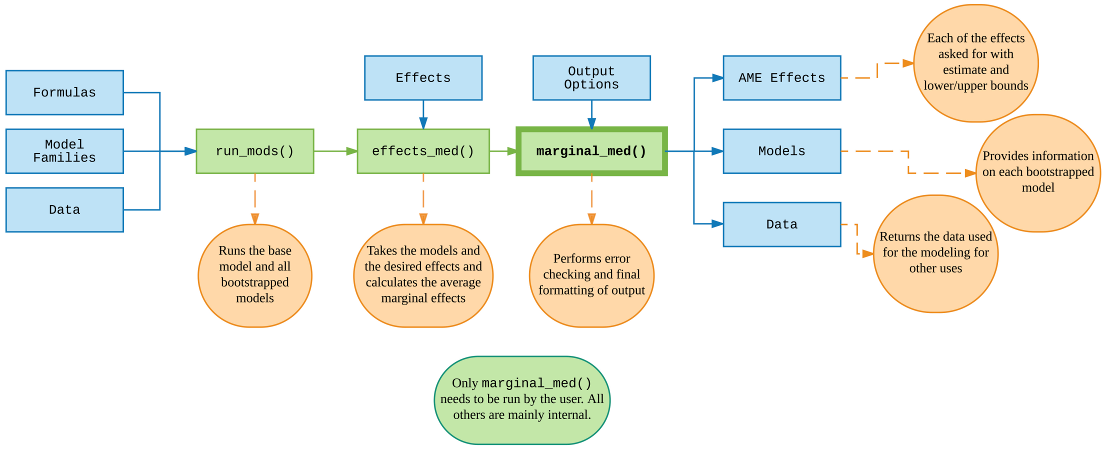

Figure 5.1: Diagram of the primary functions of the Marginal Mediation software in R.

Phase 2: Monte Carlo Simulation Study of Marginal Mediation

The evaluation of the new method is an essential step in understanding its properties and robustness and further assess the performance of the software. It is expected that Marginal Mediation will be consistent (as are average marginal effects; see Chapter 3) in estimating the underlying probabilities within each path, thus providing accurate measures for both the indirect and direct effects. This expectation will be evaluated via a Monte Carlo simulation study (Paxton, Curran, Bollen, Kirby, & Chen, 2001).

In the study, data will be simulated to come from a population of known parameters. A thorough literature review of mediation analysis in prevention work will inform on several of the facets of the study, including appropriateness of the selected sizes of the various mediated paths, the sample sizes, and the common combinations of categorical mediators and/or outcomes in these studies. Further, the insights provided in Phase I will inform on any adjustments that must be made regarding the simulation study (e.g., relax other assumptions that are not currently included in the simulation design). The results of the simulation study will help in the development of the guidelines for using Marginal Mediation in practice.

Literature Review

Before performing the Monte Carlo simulations a thorough review of the literature is necessary (Paxton et al., 2001). This review will focus on the use of mediation analysis in prevention research where the data has categorical mediators and/or outcomes.[3]

[3]Those studies found in the literature search will also be considered for replication via Marginal Mediation in later studies not part of this proposal.

Table 5.1: The various experimental conditions of the Monte Carlo simulation study.

Independent Variables	Conditions
Sample Size	200, 500
Effect size of a path - (probability units)	.05, .10, .20, .50
Effect size of b path - standardized	.10, .30, .50
Effect size of c' path - standardized	0, .30
Proportion of Positive Response (Binary Mediators)	.05, .10, .50
Number of Mediators in Model	1, 3
Relationship of Probability	Logistic, Linear
Bootstrap Size	500, 1000
Total Conditions	1,152

Simulation Studies

A Monte Carlo simulation study, via the R statistical environment version 3.4.1, will assess the finite properties of the proposed method. Monte Carlo simulation was selected due to its simplicity in generating informative results and its high success in the literature (e.g., Graham, Olchowski, & Gilreath, 2007; Nylund, Asparouhov, & Muthen, 2007). Here, 500 data sets will be simulated for each condition (Paxton et al., 2001). The data will be simulated from a known population with a researcher specified causal model (i.e., the "population model"). The model will consist of a binary mediator (0 = "No", 1 = "Yes"), a continuous outcome, a continuous predictor and a dichotomous predictor (dummy coded).

The a path population model is defined below, where the $Prob(M = 1)$ is a latent continuous variable with a linear or logistic relationship with the predictors (only linear is shown; see Table 5.1) and the ϵ_i is normally distributed with a mean of 0 and a standard deviation of 1.

$$Prob(M = 1)_i = a_0 + a_1 x_{cont} + a_2 x_{dummy} + \epsilon_i \tag{5.1}$$

The observed variable, M_i, is defined as follows:

$$M_i = \begin{cases} 0, & \text{if } Prob(M = 1) < .5, \\ 1, & \text{otherwise.} \end{cases} \tag{5.2}$$

The b and c' paths population model is identical to Equation 4.2 with only two predictors and a single binary mediator.

Table 5.1 highlights the conditions that will be varied for each simulation. A distinct Marginal Mediation model will be applied to each of the 500 data sets for each possible combination of experimental conditions. This means 576,000 total models will be fit.

The focus of the simulation study will be to gauge the accuracy of the average marginal effects at estimating the population effects while undergoing the experimental conditions. The dependent variables will be: bias (i.e., is the mean of the estimates at the population mean?), power (i.e., how often does the null properly get rejected?), confidence interval coverage (i.e., does the confidence interval cover the proper interval?), and how closely $a \times b + c'$ is to c (i.e., does the indirect plus the direct effect equal the total effect?). The effects of the conditions on these outcomes will be assessed via visualizations, descriptive tables, and regression—testing the main effects of each variable.

Guideline Development

Recommendations from the simulation study will be documented, including necessary sample sizes, bias in various conditions, and the needed number of bootstrapped samples for proper confidence intervals. The documentation will be available in manual form online on the R website, GitHub, and arXiv.org.

Potential Problems and Solutions

Two potential problems are notable. First, the vast amount of computation and data inherent in Monte Carlo studies may overwhelm the laboratory computer. If this is the case, the "super-computer" cluster available to the Prevention Science Laboratory will be used. Second, if the effect sizes, the dimensions of the data, and other nuances common in prevention data found in the literature review are vastly different than that proposed, the proposed conditions will be adjusted accordingly.

Phase 3: Application of Marginal Mediation

During the third phase, all important aspects of Marginal Mediation discovered throughout the first two phases will be used to assess the mediated effects of the relationship between chronic illnesses and delinquency/substance use among adolescents. The application will also demonstrate the ability to include latent class analysis into the framework.

Application

An application study that both explicitly demonstrates the new method and highlights new substantive work will be undertaken. The research questions inherent in this study are informed by a major meta analysis (Pinquart & Shen, 2011) and an original study (Surís, Michaud, Akre, & Sawyer, 2008). In these studies, it is clear that chronic illness—including migraines and asthma—is associated with delinquency, depression, anxiety, and substance use. These studies did not explicitly assess a mediated effect, although depression and anxiety are likely mediators of the relationship between chronic illness and delinquency/substance use (Deykin, Levy, & Wells, 1986; Manasse & Ganem, 2009).

To assess these relations, data from the National Longitudinal Study of Adolescent Health ("Add Health"; Harris, Halpern, Haberstick, & Smolen, 2013) will be used. These data include longitudinal information on adolescents (n = 14,960, at wave I), including their delinquent and substance use patterns. The measures of interest consisted of: wave I parent report of the chronic conditions ("For each of the follow-

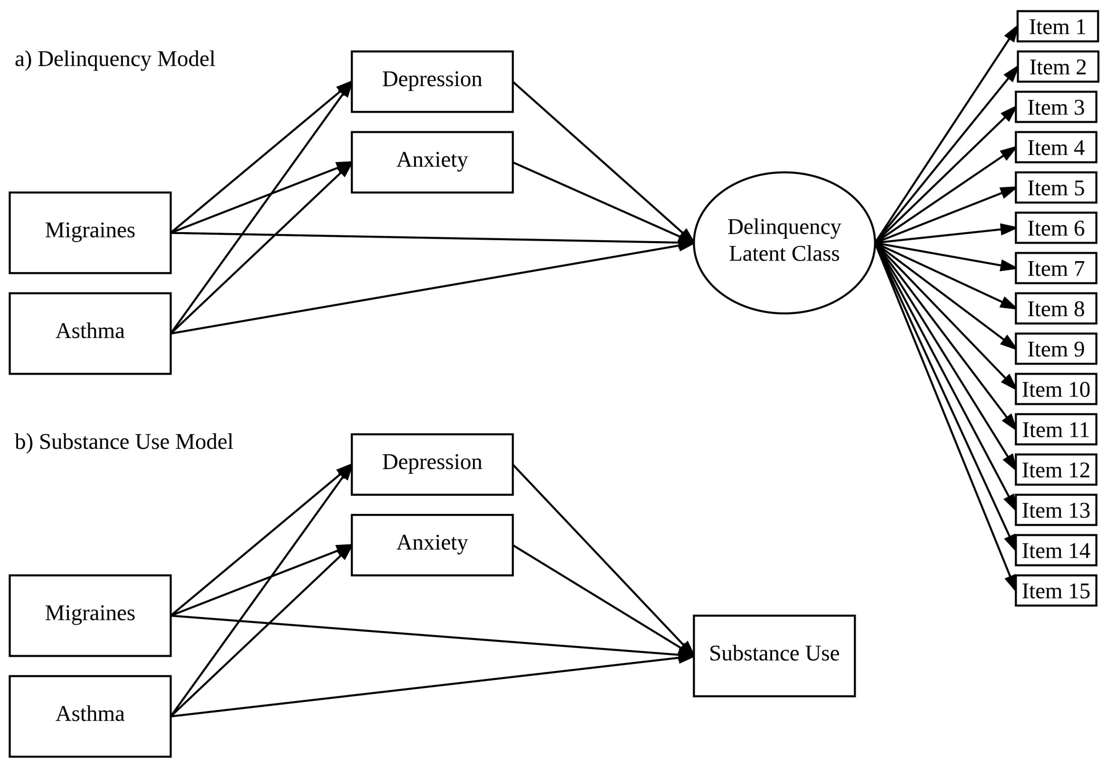

Figure 5.2: Path diagram of the two theoretical models: a) with delinquency and b) with substance use. Both outcomes are categorical, with delinquency being a latent categorical variable and substance use being observed categorical.

ing health conditions, please tell me whether [your child] has [asthma, migraines].") and wave II self-report of delinquency (consisting of 15 items) and substance use ("During the past 30 days, how many times have you used marijuana?" and "During the past two weeks, how many times did you have 5+ drinks on a single occasion?"). The mediators will be depression (19 items taken from the CES-D Scale; Radloff, 1977) and anxiety (a collection of 7 items relating to general symptoms of anxiety) from wave I. Additional variables (both covariates and mediators), as informed by the literature review, can be included as well.

Ultimately, the study will assess the path between migraines/asthma (and their interaction) with delinquency and substance use. Figure 5.2 highlights the theoretical model (not explicitly showing the interaction of migraines and asthma), where both outcomes (i.e., delinquency and substance use) are categorical; delinquency will be categorical—based on a latent class analysis—and substance use is an observed four level multinomial variable (smokes marijuana, drinks heavily, both marijuana and heavy drinking, or neither behavior).

Notably, the latent class analysis will be based on a manuscript in preparation showing several latent classes of delinquency, including a violent class, a theft-oriented class, and an all-around delinquent class. Although this was based on wave I data, similar findings are expected in wave II. Using this latent class analysis, latent class regression will be performed for all b and c' classes. The post-estimation average marginal effects will be calculated for each path. The interpretation of the effects will then inform on the difference in the risk of being in a certain delinquent class (or having used marijuana or alcohol) for adolescents that have migraines (or asthma or both) through either the indirect or the direct effect.

Potential Problems and Solutions

Some problems regarding the application study are of concern. First, although unlikely, depression and anxiety may not be important mediators (as informed by the literature review). If so, other variables may be included (e.g., family and peer relationships, other health-risk behaviors, or school performance). Additionally, the measures used in the application are self- or parent-report. There is likely measurement error, but as it is in many large surveys, there is no reliable way to remove it and will, therefore, be a limitation. Finally, the application study may bring nuances of the method to light (e.g., problems with the software). These can help inform the guidelines and can provide opportunities for further development.

Strengths and Weaknesses

The Marginal Mediation approach allows meaningful interpretation regarding mediated effects. It is expected that improving interpretation and increasing the fit of the modeling procedures with more data situations will increase both the results and inference reproducibility discussed in Chapter 1. Several facets of the project enhance the strength and increase the likelihood of success. First, the initial integration and preliminary construction of the software have taken place, providing valuable information on the planning of the method. Second, by developing the software in R, researchers around the world can apply it at no cost. Third, by assessing the finite properties via Monte Carlo simulations, there will be guidelines for its use immediately available upon dissemination. Fourth, by applying Marginal Mediation to prevention data, the walk-through of Marginal Mediation will be provided and real-world mediated effects will be explored.

However, the Marginal Mediation approach has two notable limitations. First, mediation analysis assumes no measurement error in the mediators. Although latent variable methods can help with this (Iacobucci, 2008; Lockhart et al., 2011), the data necessary are not always available. Therefore, the estimates are only as good as the data used. Second, it may also be difficult for researchers to accept given the newness of average marginal effects in the field. This will be alleviated through the use of various introductions to average marginal effects and its use in other fields.

Conclusions

The goal of this proposal is to develop, evaluate and apply a method that can provide meaningful interpretation in mediation when the mediator and/or outcome is categorical—ultimately increasing the reproducibility and application of research results. Each phase builds on this goal, providing a final product that is both in theory, software, and application useful for prevention scientists for a wide variety of research situations.

REFERENCES

Bartus, T. (2005). Estimation of marginal effects using margeff. *Stata Journal*, *5*(3), 309–329. https://doi.org/The Stata Journal

Begley, C. G., & Ioannidis, J. P. A. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, *116*(1), 116–126. https://doi.org/10.1161/CIRCRESAHA.114.303819

Catanzaro, S. J., & Laurent, J. (2004). Perceived family support, negative mood regulation expectancies, coping, and adolescent alcohol use: Evidence of mediation and moderation effects. *Addictive Behaviors*, *29*(9), 1779–1797. https://doi.org/10.1016/j.addbeh.2004.04.001

Chang, A. C., & Li, P. (2015). Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". *Finance and Economics Discussion Series*, *083*, 1–26. https://doi.org/10.17016/FEDS.2015.083

Coie, J. D., Watt, N. F., West, S. G., Hawkins, J. D., Asarnow, J. R., Markman, H. J., … Long, B. (1993). The Science of Prevention. *American Psychologist*, *48*(10), 1013–1022. https://doi.org/10.1037/0003-066X.48.10.1013

Cumming. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Deykin, E. Y., Levy, J. C., & Wells, V. (1986). Adolescent Depression, Alcohol and Drug Abuse. *American Journal of Public Health*, *77*(2), 178–182.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological Methods*, *12*(1), 1–22. https://doi.org/10.1037/1082-989X.12.1.1

Ennett, S. T., Bauman, K. E., Pemberton, M., Foshee, V. a, Chuang, Y. C., King, T. S., & Koch, G. G. (2001). Mediation in a family-directed program for prevention of adolescent tobacco and alcohol use. *Preventive Medicine*, *33*(4), 333–46. https://doi.org/10.1006/pmed.2001.0892

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science*, *10*(2), 87–99. https://doi.org/10.1007/s11121-008-0109-6

Fidler, F., Hoekstra, R., Cumming, G., Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., … Van Assen, M. A. L. M. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*(November), 1–12. https://doi.org/10.3389/fpsyg.2016.01832

Goodman, S. N. (2016). Aligning statistical and scientific reasoning. *Science*, *352*(6290), 1180–1181. https://doi.org/10.1126/science.aaf5406

Goodman, S. N., Fanelli, D., & Ioannidis, J. P. A. (2016). What does research reproducibility mean? *Science Translational Medicine*, *8*(341), 1–6. https://doi.org/10.1126/scitranslmed.aaf5027

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, *8*(3), 206–213. https://doi.org/10.1007/s11121-007-0070-9

Harris, K. M., Halpern, C. T., Haberstick, B. C., & Smolen, A. (2013). The National Longitudinal Study of Adolescent Health (Add Health) sibling pairs data. *Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies*, *16*(1), 391–8. https://doi.org/10.1017/thg.2012.137

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Statistical Mediation Analysis in the New Millennium*, *76*(4), 408–420. https://doi.org/10.1080/

03637750903310360

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. The Guilford Press.

Heuer Jr., R. J. (1999). *Psychology of intelligence analysis*. Center for the Study of Intelligence.

Hoeppner, B., Hoeppner, S., & Abroms, L. (2017). How do text-messaging smoking cessation interventions confer benefit? A multiple mediation analysis of Text2Quit. *Addiction*, 673–682. https://doi.org/10.1111/add.13685

Iacobucci, D. (2008). *Mediation analysis*. SAGE [Kindle Version].

Iacobucci, D. (2012). Mediation analysis and categorical variables: The final frontier. *Journal of Consumer Psychology*, *22*(4), 582–594. https://doi.org/10.1016/j.jcps.2012.03.006

Iacobucci, D., Saldanha, N., & Deng, X. (2007). A mediation on mediation: Evidence that structural equation models perform better than regression. *Journal of Consumer Psychological*, *7*(2), 140–154. https://doi.org/10.1016/S1057-7408(07)70020-7

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, *15*(4), 309–34. https://doi.org/10.1037/a0020761

Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, *25*(1), 51–71. https://doi.org/10.1214/10-STS321

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), 0696–0701. https://doi.org/10.1371/journal.pmed.0020124

Lockhart, G., Mackinnon, D. P., & Ohlrich, V. (2011). Mediation analysis in psychosomatic medicine research. *Psychosomatic Medicine*, *73*(1), 29–43. https://doi.org/10.1097/PSY.0b013e318200a54b.Mediation

Lockhart, G., Phillips, S., Bolland, A., Delgado, M., Tietjen, J., & Bolland, J. (2017). Prospective Relations among Low-Income African American Adolescents' Maternal Attachment Security, Self-Worth, and Risk Behaviors. *Frontiers in Psychology*, *8*(January), 1–10. https://doi.org/10.3389/fpsyg.2017.00033

Loken, E., & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, *355*(6325), 584–585. https://doi.org/10.1126/science.aal3618

Luk, J. W., Wang, J., & Simons-Morton, B. G. (2010). Bullying Victimization and Substance Use Among U.S. Adolescents: Mediation by Depression. *Prevention Science*, *11*(4), 355–359. https://doi.org/10.1007/s11121-010-0179-0

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Routledge.

MacKinnon, D. P., & Cox, M. C. (2012). Commentary on "Mediation Analysis and Categorical Variables: The Final Frontier" by Dawn Iacobucci. *Journal of Consumer Psychology*, *22*(4), 600–602. https://doi.org/10.1016/j.jcps.2012.03.009.Commentary

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. (2007). Mediation Analysis. *Annual Review of Psychology*, *58*(Hebb 1966), 593–602. https://doi.org/10.1146/annurev.psych.58.110405.085542.Mediation

Manasse, M. E., & Ganem, N. M. (2009). Victimization as a cause of delinquency: The role of depression and gender. *Journal of Criminal Justice*, *37*, 371–378.

Montreuil, B., Bendavid, Y., & Brophy, J. (2005). What is so odd about odds? *J Can Chir*, *48*(5), 400–408.

Mood, C. (2010). Logistic regression: Why we cannot do what We think we can do, and what we can do about it. *European Sociological Review*, *26*(1), 67–82. https://doi.org/10.1093/esr/jcp006

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., … Wagenmakers, E.-j. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(January), 1–9. https://doi.org/10.1038/s41562-016-0021

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indictors. *Psychometrika*, *49*(1), 115–132.

Norton, E. C. (2012). *Log Odds and Ends*. National Bureau of Economic Research. https://doi.org/10.1002/0471725331.ch16

Nylund, K. L., Asparouhov, T., & Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *15*(1), 182. https://doi.org/10.1080/10705510701575396

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. https://doi.org/10.1126/science.aac4716

Patil, P., Peng, R. D., & Leek, J. T. (2016). What Should Researchers Expect When They Replicate Studies? A Statistical View of Replicability in Psychological Science. *Perspectives on Psychological Science*, *11*(4), 539–544. https://doi.org/10.1177/1745691616646366

Paxton, P., Curran, P. J., Bollen, K. a., Kirby, J., & Chen, F. (2001). Monte Carlo Experiments: Design and Implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, *8*(2), 287–312. https://doi.org/10.1207/S15328007SEM0802_7

Pinquart, M., & Shen, Y. (2011). Behavior problems in children and adolescents with chronic physical illness: A meta-analysis. *Journal of Pediatric Psychology*, *36*(4), 375–384. https://doi.org/http://dx.doi.org/10.1093/jpepsy/jsq104

Preacher, K. J., & Kelley, K. (2011). Effect Size Measures for Mediation Models : Quantitative Strategies for Communicating Indirect Effects, *16*(2), 93–115. https://doi.org/10.1037/a0022658

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? *Nature Review Drug Discovery*, *10*, 712. https://doi.org/10.1038/nrd3439-c1

Radloff, L. (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401.

Serang, S., Jacobucci, R., Brimhall, K. C., & Grimm, K. J. (2017). Exploratory Mediation Analysis via Regularization. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–12. https://doi.org/10.1080/10705511.2017.1311775

Shih, R. A., Miles, J. N. V., Tucker, J. S., Zhou, A. J., D 'amico, E. J., Jeremy, R. O., & Miles, N. V. (2010). Racial/Ethnic Differences in Adolescent Substance Use: Mediation by Individual, Family, and School Factors*. *JOURNAL OF STUDIES ON ALCOHOL AND DRUGS / SEPTEMBER 2010 Stud. Alcohol Drugs*, *71*(September), 640–651. https://doi.org/10.15288/jsad.2010.71.640

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. *Psychological Methods*, *7*(4), 422. https://doi.org/10.1037//1082-989x.7.4.422

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Small, D. S. (2013). *Mediation Analysis Without Sequential Ignorability: Using Baseline Covariates Interacted with Random Assignment as Instrumental Variables*.

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, *13*, 290–312.

StataCorp. (2015). *Stata Statistical Software: Release 14*. StataCorp LLC.

Surís, J.-C., Michaud, P.-A., Akre, C., & Sawyer, S. M. (2008). Health risk behaviors in adolescents with chronic conditions. *Pediatrics*, *122*(5), e1113–e1118. https://doi.org/10.1542/peds.2008-1479

Taylor, J., & Tibshirani, R. J. (2015). Statistical learning and selective inference. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(25). https://doi.org/10.1073/pnas.1507583112

Wang, J., Simons-Morton, B. G., Farhart, T., & Luk, J. W. (2009). Socio-demographic variability in adolescent substance use: Mediation by parents and peers. *Prevention Science*, *10*(4), 387–396. https://doi.org/10.1007/s11121-009-0141-1

Wickham, H. (2015). *R Packages: Organize, Test, Document, and Share Your Code* (1st ed.). O'Reilly Media.